Carnegie Mellon University

Recommendation Systems in E-Commerce (H&M)

Machine Learning in Practice - 17691 Spring 22

Connie He, Phoebe Li, Huilin Xiong

Agenda

- Project Overview & Value
- Model Selection (Baseline)
- Product Launch Architecture
- Future Works
- MVP Demo

Project Overview (AI Canvas)

- Fast-Fashion Industry Dynamics. Making prediction rather than creating the trend.
- The H&M personalized fashion recommendation system will use the customer purchase data and machine learning model to make product recommendation to the customer.
- The bi-directional benefits. The recommendation system bring value to both customers and the company.

Recommendation System Architecture

- Recommends items to the customer similar to previously high-rated (frequent purchased) items by the customer.
- Uses the features and properties of the item.

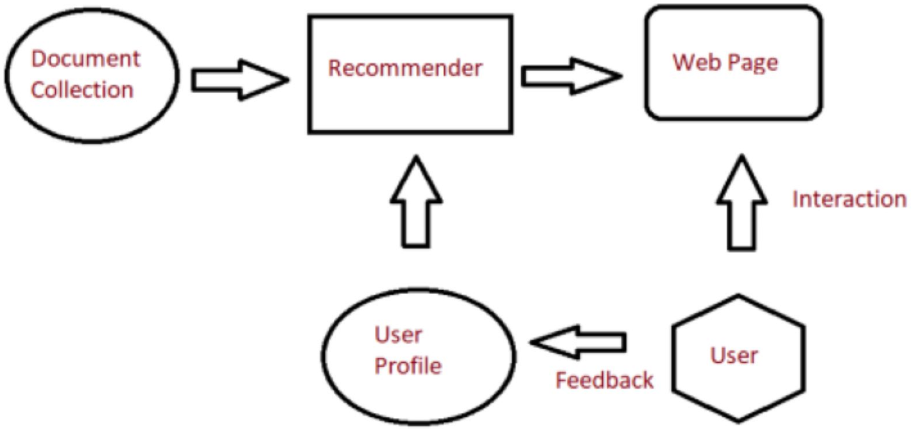
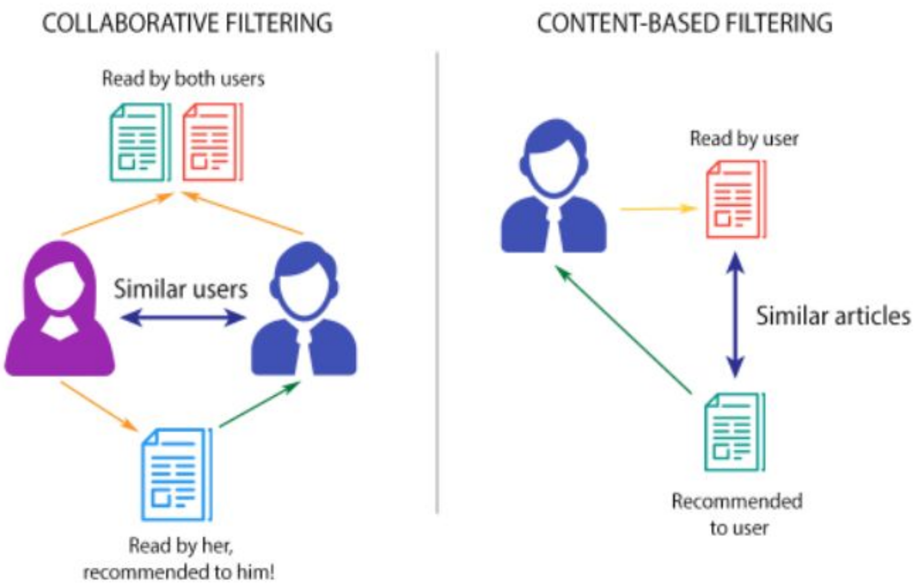


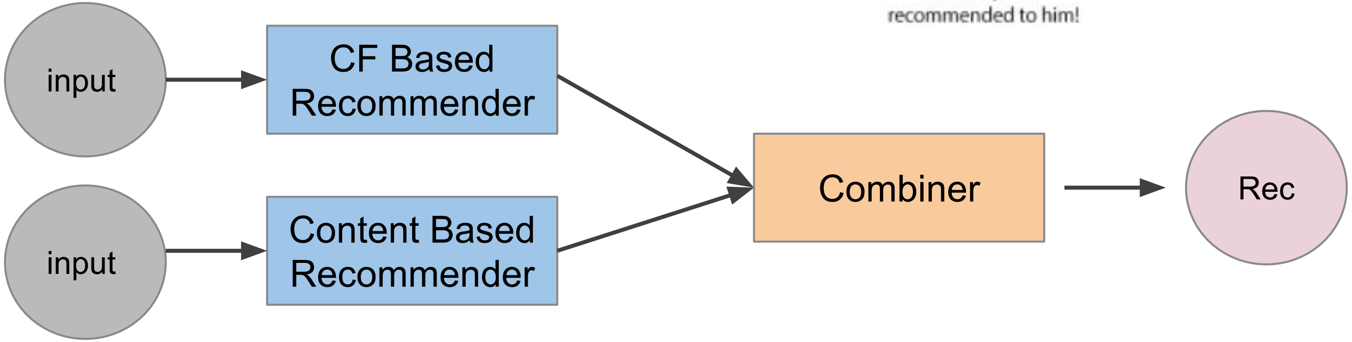
Fig: Recommender System

Source: https://www.geeksforgeeks.org/recommendation-system-in-python/

Recommendation System Overview

Collaborative filtering
Content-based filtering
Hybrid recommendation system

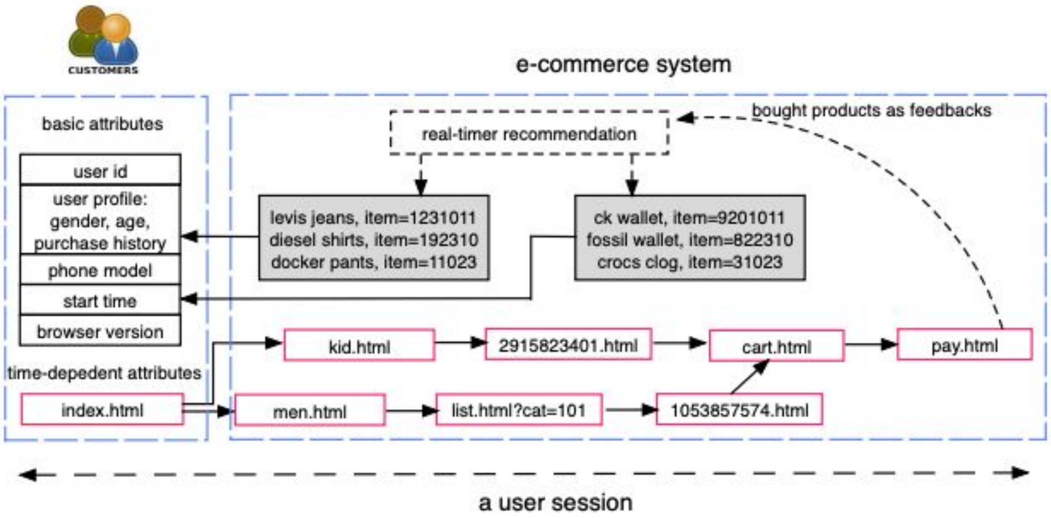


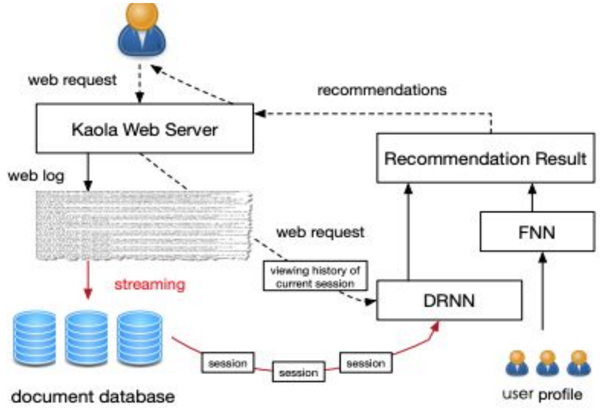


Source: https://towardsdatascience.com/what-are-product-recommendation-engines-and-the-various-versions-of-them-9dcab4ee26d5
https://towardsdatascience.com/how-can-you-tell-if-your-recommender-system-is-any-good-e4a6be02d9c2

Work Flow of Recommendation Module

Example to improve user experience, collects user's purchase history and applied DRNN (deep recurrent neural network) and FNN (feedforward neural network) using CF algorithm





Source: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7498326

Product Team

Required work roles for concept development:

Data Scientist is responsible for utilizing the transaction data and user information to train the model and provide the prediction of recommended products that the customer intends to purchase.

Data Engineer is responsible for data pipeline building and data cleaning. In the beginning, our data is gathered from Kaggle. In future, we will collect data from stores and online shopping website.

Software Engineer is responsible for develop the application or the interface for users to get the recommended products conveniently. Besides the website, we will also focus on mobile applications.

Value of Data Calculation for User

Value for **Customers:**

1. Significantly reduce the time and effort in finding the perfect match products
2. The recommendation guarantees a reliable, efficient online shopping experience

Value for **Stores:**

1. Attract more customers and more transactions and make more profits
2. Get a deep insight into the customer preferences and make more popular products
3. An innovative way to build brand loyalty
4. Increase customer retention rate

Data Network Effects

With more customers, we can gather more data and improve the model accuracy, as well as personalize the user experience. In a turn, the customers will be more willing to use the recommendation product to get their perfect match, and contribute more data for us.

When our model uses more data, we will calculate some classic user portrait, which is a classification for customers. Customers in the same classification have a high similarity in product style and preferences. Therefore, with a big data network, we can **expand our product from clothes recommendation to other lifestyle products.**

Product Feature

Reliability The data set contains 34G of real-world transactions data, guarantees the reliability of the product.

Scalability Consider hosted servers utilities (Google cloud, Azure etc.), deep learning frameworks including Tensorflow, Pytorch, MXNet etc.

Maintainability The recommendation system will be built based on the baseline model and will provide a step-by-step guideline to make it easy to maintain.

Adaptability Our product can be migrated to a website or an App with the support of cloud computing. It is adaptable to other platforms.

Model

Carnegie Mellon University

Model Selection & Metrics

- The two models based on products features and the customer purchase behaviors
- Model Selection
  - > Data Process
  - > Metrix creation
- Evaluation Metrics: recmetrics package in Python

Model Part I: Product Similarity

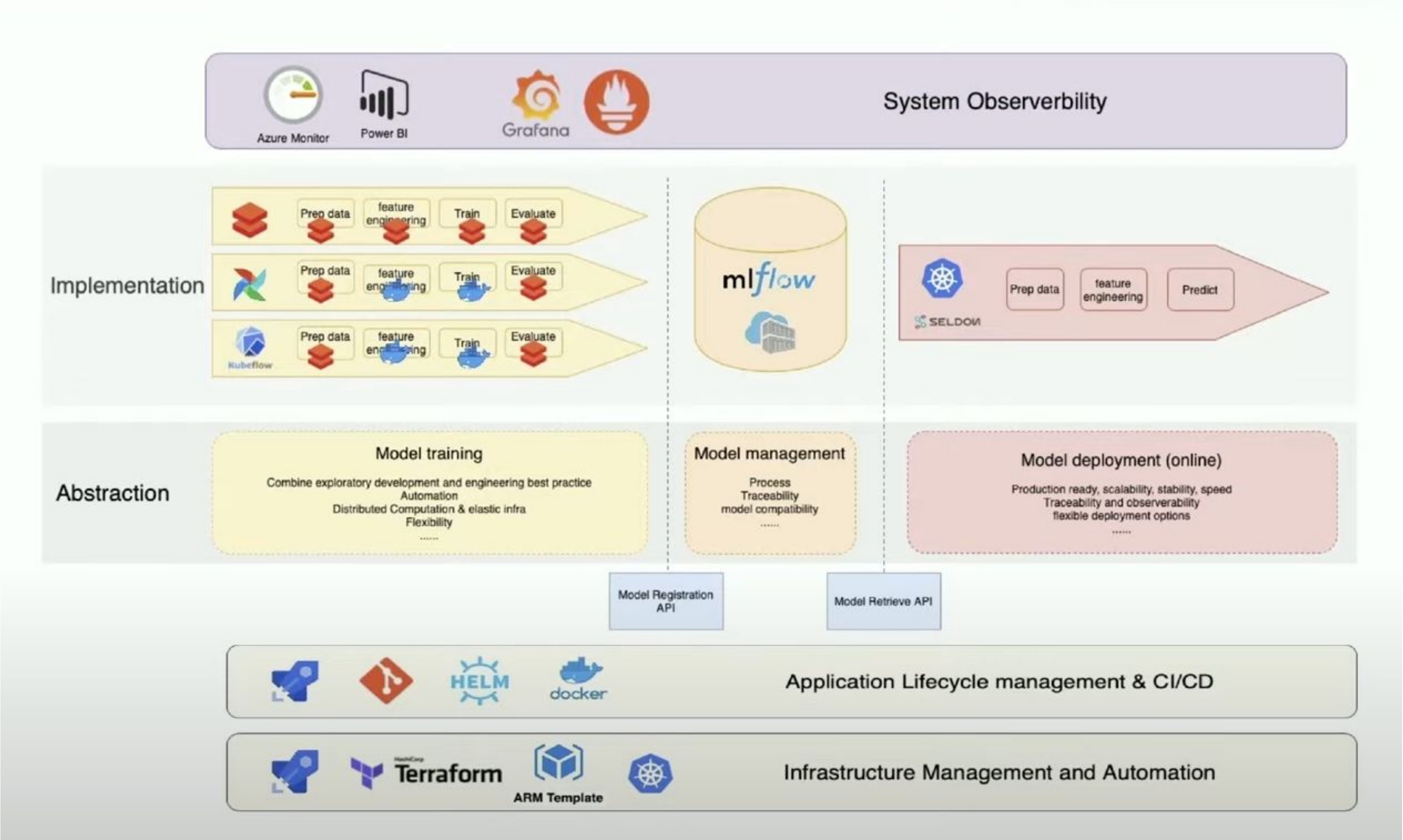
- Rearrange product information to aggregate text
- Use TFIDFVectorizor to vectorize the text data and create the tfidf_matrix
- Compute the Cosine Similarity of the Matrix
- The model will return the top 10 most similar products based on purchased item artical_id

Model Part II: Transaction Data Recommendation

- Aggregate transaction data (31 million) to select the customers who made most orders (top 200), then search for their purchased products
- Creates a highly sparse matrix between customer and products, then calculates the correlation matrix
- Shows the products that are highly correlated with the the given product; in the case that customer had 0 purchase, it automatically returns the most popular product; otherwise it returns the top 5 recommendations based on the user's latest purchase

Product Launch Architecture

Carnegie Mellon University



https://databricks.com/session_na21/scaling-ai-at-hm

Database Considerations

Graph database (Neo4J) is suitable for relationship-rich data

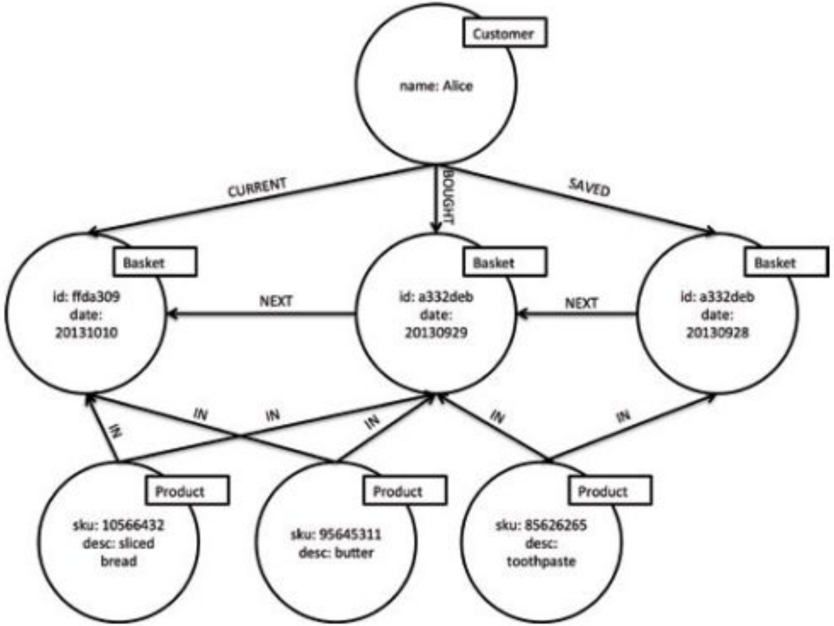
Redis to record user session, rapid access for reads and writes. No need to be durable

MongoDB to store product information, embedded nature and dynamic change

Databricks and ml flow to build the machine learning pipeline

Connects to Google machine learning ecosystem (**Google Bigtable, Bigquery**)

Neo4J Example



Source: https://neo4j.com/use-cases/real-time-recommendation-engine/

Scaling Considerations

Data Scaling

- In future, with more data collected from customers, we will only consider the previous 3-5 years purchase data.
- Besides, we will use scalable data management system to store and train the data of increasing amount, such as Apache Spark.
- Data normalization and standardization.

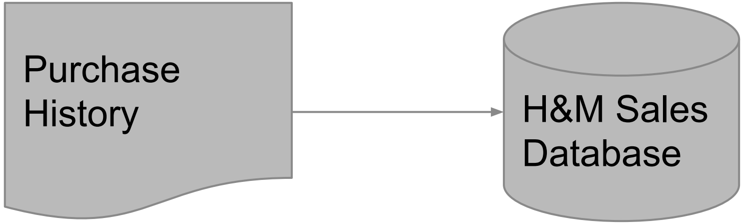
Service Scaling

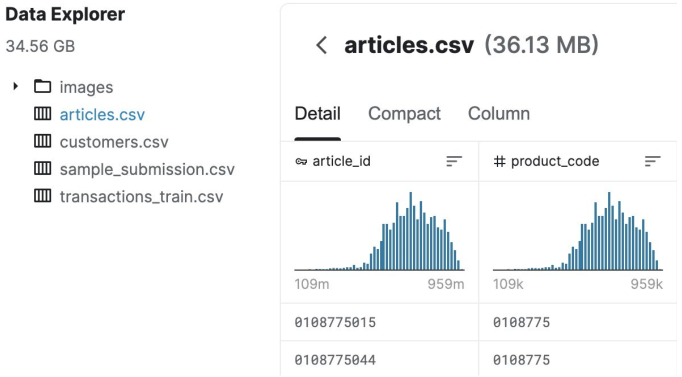
We will deploy our model on a scalable platform to handle large scale requests, such as AWS or GCP AI Platform. As traffic grows, we will add more and more VM instances and load balancer to maintain a high performance for customers.

Data Collection

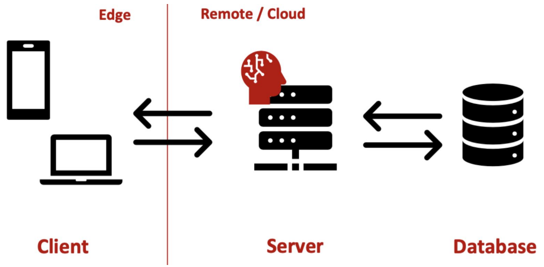
Currently, our MVP demo uses a dataset of 150,000. Kaggle provides a dataset of 34G real-world transactions data in a 7-day period. In future, we will use the latest data from shopping stores.

- Purchase History
- Basic Customer information
- Product information
- Purchasing trend





Model Deployment



- **Model as a Service**

  To guarantee the reliability and scalability, we will deploy our product on AWS EC2 Service and make full use of its elasticity to handle different amounts of requests at different time.

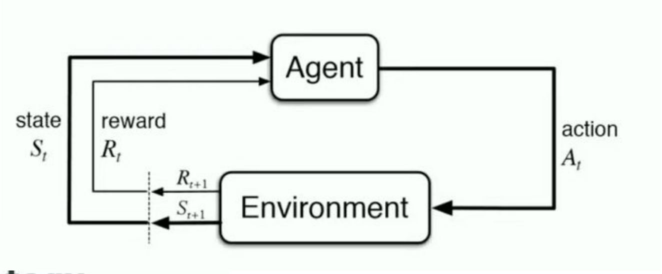
- **Deploying at the Edge**

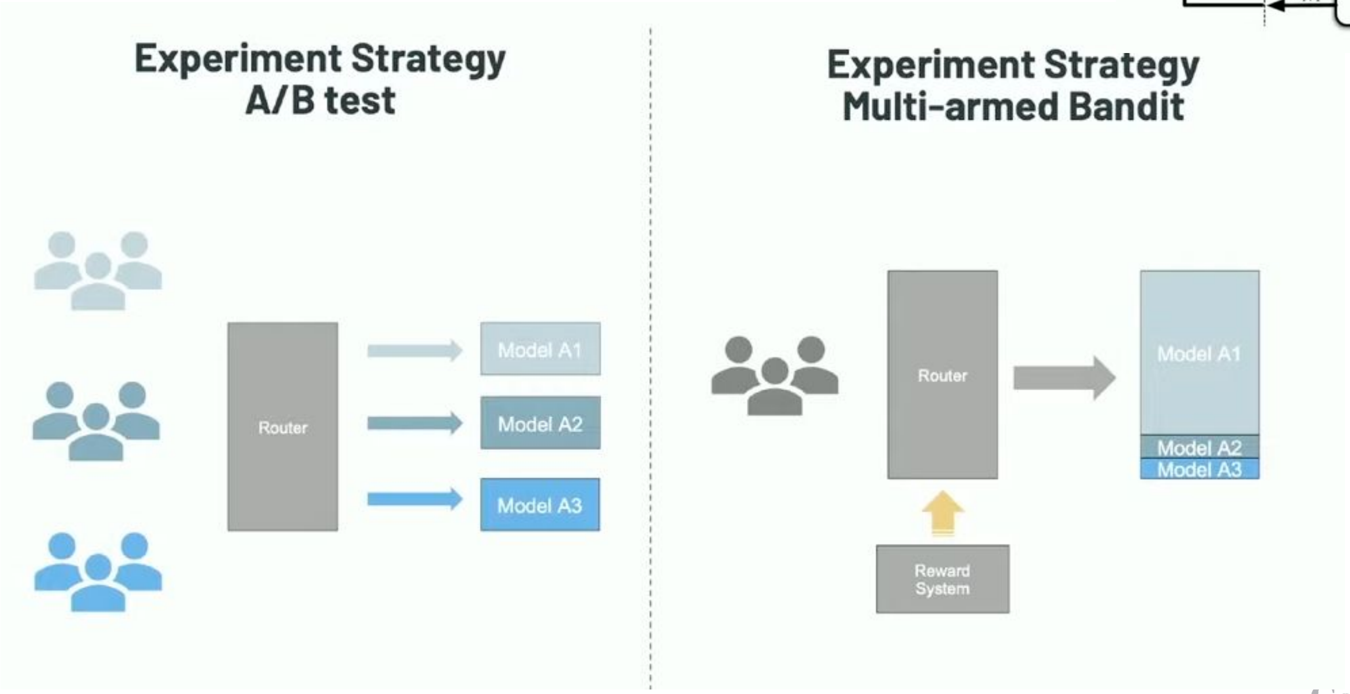
  We will use CoreML to integrate machine learning model into mobile application. So that our users can get a convenient and fast service everywhere.

Monitoring

- Continue to improve by adding extra customer browsing behaviors, product features, transaction information
- **Overall Evaluation Criterion (OEC)**
  - Determine whether the particular treatment is successful or not
  - Reflects long run objectives:
    - higher revenue, higher customer retention rate, user's lifetime values
  - Short term objectives:
    - Higher visit frequency, longer visiting sessions, disclosure of personal information, user engagement, interaction with the site, user referral rate, add to cart rate
- Measure of "**randomness**"
  - When do we conclude the results don't generate by randomness?
- **Seasonality Effect**

Testing (A/B Test)

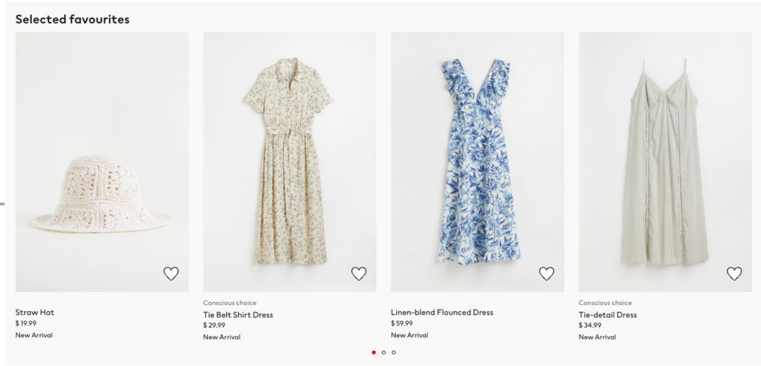


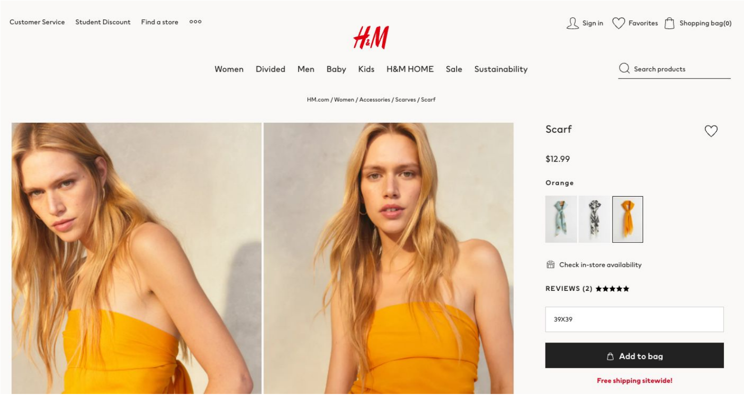


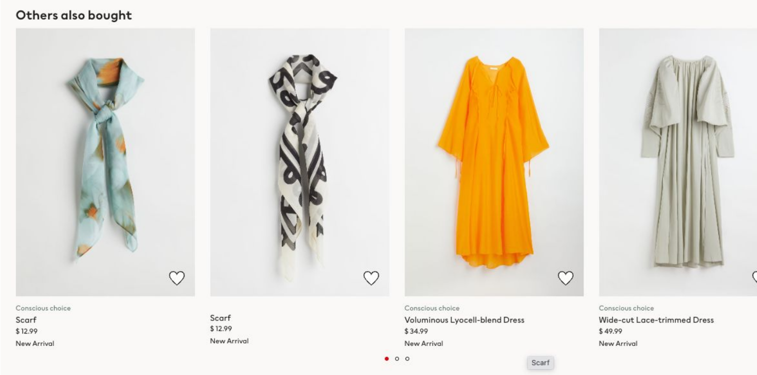
End User Application

Displays at current H&M's website:

- "**Selected favourites**" and "**Others also bought**"







Ethical Considerations

Collection and use of user data

We will specify how users' data might be used in license agreements and get a consensus with the shopping stores.

Black-box Algorithm

Our recommendation is just a product reference. If users do not like the recommended products, they can mark it as 'unlike'. This feature avoids the harm of incorrect results. In other fields like healthcare, they need to be more cautious about the result.

Future Works

Carnegie Mellon University

Model Improvements

Model:

- Feature extraction
- LSTM (long short term memory network)
- Reinforcement Learning (reward indicates client's next purchase)
- Handling product image data (CNN)

Model Retraining, **Data Version Control**

Recommendation in **Real Time**

Deployment (A/B Test)

Long Term Business Impact

How do we switch from Kaggle to real world data?

- Suppose we work at H&M and we are tasked with recommending the next product to customers
- We want to:
  - Reduce time to introduce new products to the market
  - Learning feedback loop by democratizing AI in organization
  - Unlock flow by standardizing and simplifying
- We can do:
  - **Recommendation engines**
  - **Content personalization**
  - **Customer value**
- We need to implement:
  - Dev test, stage test, deploy to production
  - **Model tracking (mlflow)**
  - **Model registry (Azure DevOps)**
- We can try:
  - Web-scrape other competitor's website information
  - Browsing behavior by cookies

Ethics

How to achieve fairness?

- Understand pre-processing data
- Use additional constraints during training
- Post-processing predictions

- **Independence**: $h(\vec{x}, a) \perp a$
  - Probability of being accepted is the same for all genders
- **Separation**: $h(\vec{x}, a) \perp a \mid y$
  - All "good" applicants are accepted with the same probability, regardless of gender
  - Same for all "bad" applicants
- **Sufficiency**: $y \perp a \mid h(\vec{x}, a)$
  - For the purposes of predicting y, the information contained in $h(\vec{x}, a)$ is "sufficient", a becomes irrelevant
- Any two of these criteria are mutually exclusive in the general case!

Source:
http://www.cs.cmu.edu/~mgormley/courses/10601/slides/lecture27-final.pdf

Demo

MVP Demo

Demo Colab

Carnegie Mellon University

References

Databricks, H&M Group: Revolutionizing Fashion with AI, Retrieved April 18, 2022, from https://databricks.com/customers/hm

Databricks, (2022, March 8). Collaborating Across Retail Value Chain With Data and AI, Retrieved April 19, 2022, from https://databricks.com/explore/retail-resources/how-a-retail-lakehouse-can-evolve-your-organization

Hertzberg, B, Data + AI Summit 2022, Scaling AI At H\&M, Retrieved April 17, 2022, from https://databricks.com/session_na21/scaling-ai-at-hm

Kim, C., Google Cloud, Moloco Handles 5 Million + ads requests per second with Cloud Bigtable, Retrived April 20, 2022, from https://cloud.google.com/blog/products/databases/moloco-uses-cloud-bigtable-to-build-their-ad-tech-platform